Approximating the Permanent of Positive Semidefinite Matrices

Nima Anari

Joint work with

Leonid Gurvits

Shayan Oveis Gharan

Amin Saberi

Determinant

$$\det(M) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) M_{1,\sigma(1)} \dots M_{n,\sigma(n)}$$

Permanent

$$\operatorname{per}(M) = \sum_{\sigma \in S_n} M_{1,\sigma(1)} \dots M_{n,\sigma(n)}$$

Determinant

$$\det(M) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) M_{1,\sigma(1)} \dots M_{n,\sigma(n)}$$

Permanent

$$\mathrm{per}(M) = \sum_{\sigma \in S_n} M_{1,\sigma(1)} \dots M_{n,\sigma(n)}$$

2×2 Example

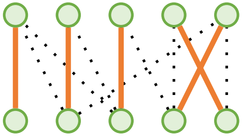
$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\det(M) = ad - bc$$

$$\mathrm{per}(M) = ad + bc$$

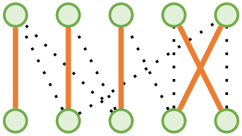
Complexity of Permanent

- #P-hard to compute $\mathrm{per}(M)$ for $0/1$ matrices [Valiant'79].



Complexity of Permanent

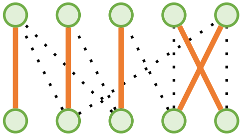
- #P-hard to compute $\mathrm{per}(\mathsf{M})$ for $0/1$ matrices [Valiant'79].



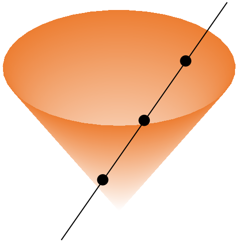
- #P-hard to compute sign of $\mathrm{per}(\mathsf{M})$ [Aaronson'11].

Complexity of Permanent

- #P-hard to compute $\mathrm{per}(M)$ for $0/1$ matrices [Valiant'79].



- #P-hard to compute sign of $\mathrm{per}(M)$ [Aaronson'11].
- #P-hard to compute $\mathrm{per}(M)$ for $M \succeq 0$ [Grier-Schaeffer'16].



Approximating the Permanent

Additive $\pm\epsilon \|M\|^n$ approximation [Gurvits'05].

Approximating the Permanent

Additive $\pm\epsilon \, |M|^n$ approximation [Gurvits'05].

Positive Matrices ($M \geq 0$)

- Permanent is always nonnegative:

$$\mathrm{per}(M) \geq 0.$$

Approximating the Permanent

Additive $\pm\epsilon\, |M|^n$ approximation [Gurvits'05].

Positive Matrices ($M \geq 0$)

- Permanent is always nonnegative:

$$\mathrm{per}(M) \geq 0.$$

- Randomized $(1+\epsilon)$-approximation (FRPAS) [Jerrum-Sinclair-Vigoda'04].

Approximating the Permanent

Additive $\pm\epsilon \, |\mathsf{M}|^{\mathsf{n}}$ approximation [Gurvits'05].

Positive Matrices ($\mathsf{M} \geq 0$)

- Permanent is always nonnegative:

$$\mathrm{per}(\mathsf{M}) \geq 0.$$

- Randomized $(1+\epsilon)$-approximation (FRPAS) [Jerrum-Sinclair-Vigoda'04].
- Deterministic 2^{n}-approximation [Gurvits-Samorodnitsky'14].

PSD Matrices ($\mathsf{M} \succeq 0$)

- Permanent is always nonnegative:

$$\mathrm{per}(\mathsf{M}) \geq 0.$$

Approximating the Permanent

Additive $\pm\epsilon\,|M|^n$ approximation [Gurvits'05].

Positive Matrices ($M \geq 0$)

- Permanent is always nonnegative:

$$\mathrm{per}(M) \geq 0.$$

- Randomized $(1+\epsilon)$-approximation (FRPAS) [Jerrum-Sinclair-Vigoda'04].
- Deterministic 2^n-approximation [Gurvits-Samorodnitsky'14].

PSD Matrices ($M \succeq 0$)

- Permanent is always nonnegative:

$$\mathrm{per}(M) \geq 0.$$

- Deterministic $n!$-approximation [Marcus'63]: $M_{1,1} \dots M_{n,n}$.

Approximating the Permanent

Additive $\pm\epsilon\,|M|^n$ approximation [Gurvits'05].

Positive Matrices ($M \geq 0$)

- Permanent is always nonnegative:

$$\mathrm{per}(M) \geq 0.$$

- Randomized $(1+\epsilon)$-approximation (FRPAS) [Jerrum-Sinclair-Vigoda'04].
- Deterministic 2^n-approximation [Gurvits-Samorodnitsky'14].

PSD Matrices ($M \succeq 0$)

- Permanent is always nonnegative:

$$\mathrm{per}(M) \geq 0.$$

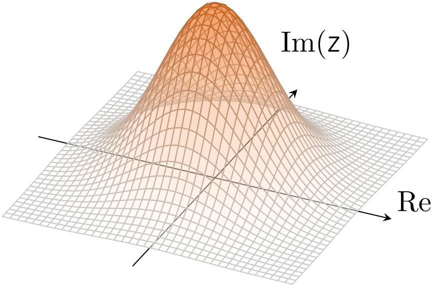
- Deterministic $n!$-approximation [Marcus'63]: $M_{1,1} \ldots M_{n,n}$.
- Improved to $\frac{n!}{k!^{n/k}}$-approximation in time $2^{O(k+\log(n))}$ [Lieb'66].

Theorem [A-Gurvits-Oveis Gharan-Saberi'17]

The permanent of PSD matrices $M \in \mathbb{C}^{n\times n}$ can be approximated, in deterministic polynomial time, within

$$(e^{\gamma+1})^n \simeq 4.84^n.$$

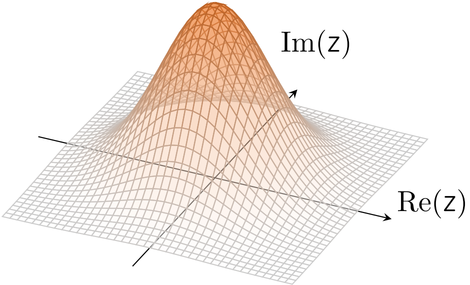
Complex Gaussians



$$z \sim \mathbb{C}\mathcal{N}(0,1)$$

$$\mathbb{P}[z] = \frac{1}{\pi} e^{-|z|^2}$$

Complex Gaussians

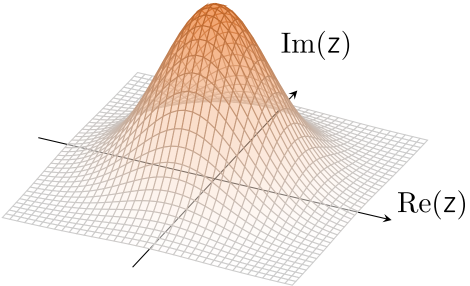


$$z \sim \mathbb{C}\mathcal{N}(0,1)$$

$$\mathbb{P}[z] = \frac{1}{\pi} e^{-|z|^2}$$

- Standard multivariate complex Gaussian: $z = (z_1, \ldots, z_n)$ i.i.d. and $z_i \sim \mathbb{C}\mathcal{N}(0,1)$.

Complex Gaussians



$$z \sim \mathbb{C}\mathcal{N}(0, 1)$$

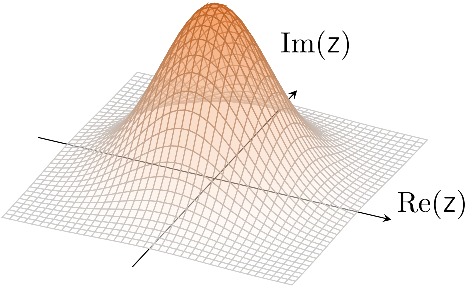
$$\mathbb{P}[z] = \frac{1}{\pi} e^{-|z|^2}$$

- Standard multivariate complex Gaussian: $z = (z_1, \ldots, z_n)$ i.i.d. and $z_i \sim \mathbb{C}\mathcal{N}(0, 1)$.
- General (circularly-symmetric) complex Gaussian:

$$g = Cz,$$

$$g \sim \mathbb{C}\mathcal{N}(0, CC^\dagger).$$

Complex Gaussians



$$z \sim \mathbb{C}\mathcal{N}(0,1)$$

$$\mathbb{P}[z] = \frac{1}{\pi} e^{-|z|^2}$$

- Standard multivariate complex Gaussian: $z = (z_1, \ldots, z_n)$ i.i.d. and $z_i \sim \mathbb{C}\mathcal{N}(0,1)$.
- General (circularly-symmetric) complex Gaussian:

$$g = Cz,$$

$$g \sim \mathbb{C}\mathcal{N}(0, CC^\dagger).$$

Wick's Formula

$$\mathbb{E}\left[|g_1|^2 \ldots |g_n|^2\right] = \mathrm{per}(CC^\dagger).$$

Schur Power

The Schur power of an $n \times n$ matrix M is

$$n! \left\{ \underbrace{\begin{bmatrix} \vdots & \vdots & \vdots \\ \vdots & M_{\sigma(1),\tau(1)} \cdots M_{\sigma(n),\tau(n)} & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}}_{n!} \right.$$

Schur Power

The Schur power of an $n \times n$ matrix M is

$$n! \left\{ \underbrace{\begin{bmatrix} \vdots & \vdots & \vdots \\ \vdots & M_{\sigma(1),\tau(1)} \cdots M_{\sigma(n),\tau(n)} & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}}_{n!} \right.$$

- The Schur power is a minor of $M^{\otimes n}$.

$$M \succeq 0 \implies \mathrm{schur}(M) \succeq 0$$

Schur Power

The Schur power of an $n \times n$ matrix M is

$$n! \left\{ \begin{bmatrix} \vdots & & \vdots & & \vdots \\ \vdots & & M_{\sigma(1),\tau(1)} \cdots M_{\sigma(n),\tau(n)} & & \vdots \\ \vdots & & \vdots & & \vdots \end{bmatrix} \right. \underbrace{\hphantom{xxxxxxxxxxxxxxxxxxxx}}_{n!}$$

- The Schur power is a minor of $M^{\otimes n}$.

$$M \succeq 0 \implies \mathrm{schur}(M) \succeq 0$$

- The permanent is an eigenvalue:

$$\mathrm{schur}(M)1 = \mathrm{per}(M)1.$$

$$M \succeq 0 \implies \mathrm{per}(M) \geq 0$$

Schur Power

The Schur power of an $n \times n$ matrix M is

$$n! \left\{ \underbrace{\begin{bmatrix} \vdots & \vdots & \vdots \\ \vdots & M_{\sigma(1),\tau(1)} \cdots M_{\sigma(n),\tau(n)} & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}}_{n!} \right.$$

- The Schur power is a minor of $M^{\otimes n}$.

$$M \succeq 0 \implies \mathrm{schur}(M) \succeq 0$$

- The permanent is an eigenvalue:

$$\mathrm{schur}(M)\mathbb{1} = \mathrm{per}(M)\mathbb{1}.$$

$$M \succeq 0 \implies \mathrm{per}(M) \geq 0$$

- Permanent is monotone w.r.t. $\succeq$:

Permanent is Loewner-Monotone

$$M_1 \succeq M_2 \succeq 0 \implies \mathrm{per}(M_1) \geq \mathrm{per}(M_2) \geq 0$$

Approximation using Monotonicity

- Permanent is monotone w.r.t. $\succeq$:

$$\mathsf{D} \succeq \mathsf{M} \succeq \mathsf{v}\mathsf{v}^\dagger \implies \mathrm{per}(\mathsf{D}) \geq \mathrm{per}(\mathsf{M}) \geq \mathrm{per}(\mathsf{v}\mathsf{v}^\dagger).$$

Approximation using Monotonicity

- Permanent is monotone w.r.t. $\succeq$:

$$D \succeq M \succeq vv^\dagger \implies \mathrm{per}(D) \geq \mathrm{per}(M) \geq \mathrm{per}(vv^\dagger).$$

Theorem [A-Gurvits-Oveis Gharan-Saberi'17]

For any $M \succeq 0$ there exist diagonal matrix D and rank-1 matrix $vv^\dagger$ such that

$$D \succeq M \succeq vv^\dagger,$$

and $\mathrm{per}(D) \leq 4.85^n \, \mathrm{per}(vv^\dagger)$.

Computing the Approximation

- Solve and output the following

$$
\begin{array}{ll}
\inf_{\mathsf{D}} & \operatorname{per}(\mathsf{D}), \\
\text{subject to} & \mathsf{D} \succeq \mathsf{M}.
\end{array}
$$

Computing the Approximation

- Solve and output the following

$$\begin{array}{ll} \inf_{\mathsf{D}} & \mathrm{per}(\mathsf{D}), \\ \text{subject to} & \mathsf{D} \succeq \mathsf{M}. \end{array}$$

- Equivalently solve the convex program

$$\begin{array}{ll} \inf_{\mathsf{D}^{-1}} & \log(\mathrm{per}((\mathsf{D}^{-1})^{-1}), \\ \text{subject to} & \mathsf{M}^{-1} \succeq \mathsf{D}^{-1} \succeq 0. \end{array}$$

- No such convex program for the best rank-1 matrix.

Sketch of Proof

- Renormalize rows and columns to assume $D = I$.

Sketch of Proof

- Renormalize rows and columns to assume $D = I$.
- By duality, there is $B \succeq 0$ with $\mathrm{diag}(B) = 1$ such that $(I - M)B = 0$:

$$B = MB.$$

  B is called a correlation matrix.

Sketch of Proof

- Renormalize rows and columns to assume $D = I$.
- By duality, there is $B \succeq 0$ with $\mathrm{diag}(B) = 1$ such that $(I - M)B = 0$:

$$B = MB.$$

  B is called a correlation matrix.
- Let $P = \mathrm{proj}_{\mathrm{imag}(B)}$. Then $M \succeq P$ because

$$x \in \mathrm{imag}(B) \implies x = By \implies Mx = MBy = By = x = Px.$$

Sketch of Proof

- Renormalize rows and columns to assume $D = I$.
- By duality, there is $B \succeq 0$ with $\mathrm{diag}(B) = 1$ such that $(I - M)B = 0$:

$$B = MB.$$

  B is called a correlation matrix.
- Let $P = \mathrm{proj}_{\mathrm{imag}(B)}$. Then $M \succeq P$ because

$$x \in \mathrm{imag}(B) \implies x = By \implies Mx = MBy = By = x = Px.$$

- Prove the "PSD Van der Waerden"

PSD Van der Waerden [A-Gurvits-Oveis Gharan-Saberi'17]

If B is a correlation matrix and P the orthogonal projection onto the image of B, then

$$\mathrm{per}(P) \geq 4.85^{-n}.$$

PSD Van der Waerden

- Given correlation matrix B (i.e. $\mathsf{B} \succeq 0$ and $\mathrm{diag}(\mathsf{B}) = 1$), want to show

$$\mathrm{per}(\mathrm{proj}_{\mathrm{imag}(\mathsf{B})}) \geq 4.85^{-\mathsf{n}}.$$

PSD Van der Waerden

- Given correlation matrix B (i.e. $\mathsf{B} \succeq 0$ and $\mathrm{diag}(\mathsf{B}) = 1$), want to show

$$\mathrm{per}(\mathrm{proj}_{\mathrm{imag}(\mathsf{B})}) \geq 4.85^{-\mathsf{n}}.$$

- Show for some unit vector $\mathsf{v} \in \mathrm{imag}(\mathsf{B})$

$$\mathrm{per}(\mathsf{v}\mathsf{v}^{\dagger}) \geq 4.85^{-\mathsf{n}}.$$

PSD Van der Waerden

- Given correlation matrix B (i.e. $B \succeq 0$ and $\mathrm{diag}(B) = 1$), want to show

$$\mathrm{per}(\mathrm{proj}_{\mathrm{imag}(B)}) \geq 4.85^{-n}.$$

- Show for some unit vector $v \in \mathrm{imag}(B)$

$$\mathrm{per}(vv^\dagger) \geq 4.85^{-n}.$$

- Let B be the Gram matrix of unit vectors $u_1, \dots, u_n$. Generate v by normalizing the projection vector of $u_1, \dots, u_n$ onto some direction g

$$v = \frac{[g^\dagger u_1 \dots g^\dagger u_n]}{|[g^\dagger u_1 \dots g^\dagger u_n]|}.$$

GM-AM Ratio

- Let u be a random vector (e.g., uniformly sampled from $u_1, \ldots, u_n$). Define the GM-AM ratio as:

$$\frac{e^{\mathbb{E}\left[\log(|u|^2)\right]}}{\mathbb{E}\left[|u|^2\right]}$$

GM-AM Ratio

- Let u be a random vector (e.g., uniformly sampled from $u_1, \ldots, u_n$). Define the GM-AM ratio as:

$$\frac{e^{\mathbb{E}\left[\log(|u|^2)\right]}}{\mathbb{E}\left[|u|^2\right]}$$

- The GM-AM ratio is always ≤ 1. Equality happens when $|u| = 1$.

GM-AM Ratio

- Let u be a random vector (e.g., uniformly sampled from $u_1, \ldots, u_n$). Define the GM-AM ratio as:

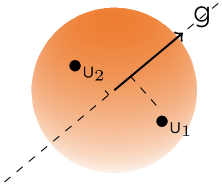
$$\frac{e^{\mathbb{E}\left[\log(|u|^2)\right]}}{\mathbb{E}\left[|u|^2\right]}$$

- The GM-AM ratio is always ≤ 1. Equality happens when $|u| = 1$.

Lemma [A-Gurvits-Oveis Gharan-Saberi'17]

If u is a random unit vector, there exists g such that the GM-AM ratio of $g^\dagger u$ is at least

$$e^{-\gamma}.$$



Complex Gaussians Come Back

- Let g be a standard complex Gaussian. Then with positive probability we have:

$$\text{GM-AM}(g^\dagger u) \geq \frac{\mathbb{E}\left[e^{\mathbb{E}\left[\log(|g^\dagger u|^2)\right]}\right]}{\mathbb{E}\left[|g^\dagger u|^2\right]} \geq \frac{e^{\mathbb{E}\left[\log(|g^\dagger u|^2)\right]}}{\mathbb{E}\left[|g^\dagger u|^2\right]}$$

Complex Gaussians Come Back

- Let g be a standard complex Gaussian. Then with positive probability we have:

$$\text{GM-AM}(g^\dagger u) \geq \frac{\mathbb{E}\left[e^{\mathbb{E}\left[\log(\left|g^\dagger u\right|^2)\right]}\right]}{\mathbb{E}\left[|g^\dagger u|^2\right]} \geq \frac{e^{\mathbb{E}\left[\log(\left|g^\dagger u\right|^2)\right]}}{\mathbb{E}\left[|g^\dagger u|^2\right]}$$

- But

$$\mathbb{E}\left[\log(\left|g^\dagger u\right|^2)\right] = -\gamma,$$

and

$$\mathbb{E}\left[\left|g^\dagger u\right|^2\right] = 1.$$

Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.

Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.
- Analysis is tight. Can we improve by sandwiching between block-diagonal matrices and higher rank matrices?

Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.
- Analysis is tight. Can we improve by sandwiching between block-diagonal matrices and higher rank matrices?
- Use lifts to get better approximations?

Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.
- Analysis is tight. Can we improve by sandwiching between block-diagonal matrices and higher rank matrices?
- Use lifts to get better approximations?
- Markov chains to get $(1+\epsilon)$-approximation?

Conclusion and Open Questions

- $(e^{\gamma+1})^n$-approximation for the permanent of PSD matrices.
- Analysis is tight. Can we improve by sandwiching between block-diagonal matrices and higher rank matrices?
- Use lifts to get better approximations?
- Markov chains to get $(1+\epsilon)$-approximation?

thank you!